Equation of State Fitting with the REFPROP Program

Eric W. Lemmon

Physical and Chemical Properties Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305

January 5, 2010

Information below in red indicates material added or changed recently.

Introduction

Fitting equations of state to represent the thermodynamic properties of pure fluids and mixtures is an art form that requires a substantial amount of background knowledge to obtain results that not only fit the experimental data, but that extrapolate well and behave appropriately as defined by a number of constraints. Unlike simple, linear, two dimensional x-y fitting procedures, fitting equations of state requires three dimensional thinking of first, second, and third derivatives of a fundamental surface (the Helmholtz energy) that is not intuitively visualized. Further insight is obtained with fourth and fifth derivatives of the surface, which can make it possible to fit good equations with very limited data. Because there is so much more involved with developing equations of state, the correlator becomes more of an artist than a statistician, and he or she should be more focused on the behavior, extrapolation, and deviations of the equation rather than on simple global statistics. The fitting described here is not so much about obtaining the lowest AAD in any particular data set, but more about finding a solution that obtains the best compromise among different data types and that demonstrates the best thermodynamic consistency.

The procedure outlined here for fitting equations of state requires 1) setting up the fitting programs that use the Refprop Fortran routines, 2) doing a literature search for experimental data and entering those data into files that can be read by the fitter, 3) converting the data to consistent units and making the input files that will be used in fitting, 4) running the fit routines to develop the equation of state, 5) validating the new equation against data and extrapolation requirements, 6) developing an ancillary equation for the vapor pressure, and 7) publishing the equation.

On your first pass through this document and setting up the fitting routines, read only the information below in blue. This will get you up and running quickly without being overwhelmed with all the little details. Once you have your fit going, you can come back and read all the additional details that are not highlighted.

Setup of computer files

All of the examples given here assume that the REFPROP routines are stored on the D drive under a REFPROP directory. Inside of this directory are subdirectories labeled FLUIDS, FITTER, and FORTRAN. The use of the D drive for storing ones information eliminates the need to keep separate Window's files from user's files. Thus, a total crash of the C drive would not lose any vital information and the user should only be concerned about backing up the D drive. Other users may wish to install the files on their C drive, in which case the examples given here will need to be changed from D: to C:.

Other useful files are contained in the BIBLIO and SOURCEB directories. The BIBLIO directory contains the NIST and CATS (University of Idaho) bibliographies. These two sets are stored in subdirectories called NIST and CATS. The SOURCEB directory

contains many useful programs written in QuickBasic to do data conversion, align periods in files, calculate statistics, and so forth.

The fluid R-161 (ethyl fluoride) is used here as an example for fitting. All of the experimental data are contained in files in an R161 directory, along with a subdirectory labeled fit that gives an example for fitting an equation of state to this fluid.

The zipped files contained in the FITTER, BIBLIO, SOURCEB, and R161 directories can be found at:

http://www.boulder.nist.gov/div838/theory/refprop/REFPROP.ZIP
http://www.boulder.nist.gov/div838/theory/refprop/FITTER.ZIP
http://www.boulder.nist.gov/div838/theory/refprop/BIBLIO.ZIP
http://www.boulder.nist.gov/div838/theory/refprop/SOURCEB.ZIP
http://www.boulder.nist.gov/div838/theory/refprop/R161.ZIP

Bibliography

The files in the BIBLIO directory contain a summary of all the papers collected in the Fluids group at NIST and CATS. There are four main programs that allow you to search the bibliography in different ways. The program SEARCH.EXE will search through all of the papers in the database and display those matching the input criteria. On the first line enter the name of the file to which the output will be saved (such as c:\output.txt). On the second line you can limit what is searched by entering a code. These codes include "A" for the author string only, "T" for the titles, "J" for the journal name, "F" for the fluid name, and "P" for the property. Only one code can be entered on this line. In general you should always leave this line blank for maximum results. The third line specifies the type of input. The default is "AND" so that the search must contain all of the words you select. To select "AND", leave this line blank. The command "OR" can also be used, and the command "XAND" specifies that the first word and any of the following words must exist in the document. Thus to see all the publications by Wagner done between 2002 and 2004, one would enter "XAND" on the third line, "Wagner" on the line labeled "1:", and "2002", "2003", and "2004" on the following lines. To exclude a word, use a "~" before an entry, e.g., adding "~nitrogen" will exclude all papers dealing with nitrogen in the previous example.

The program "SRCH_ATH.EXE" allows you to enter the last name of the first author of a paper to see all papers written by that author. After pressing Enter, a list will appear and you can scroll up or down through this list. At the bottom of the screen each corresponding citation will appear as you scroll through the list. Pressing Tab will take you back to the input line to enter a new author name. Pressing F2 will print the citation to a file called "BIB.OUT". The "SRCH_NUM.EXE" program allows you to enter the assigned number for a specific paper to view the citation. The program "MAKEREF" is explained in a later section.

Conversion of experimental data

Experimental data should be entered into an ASCII file named with the first letters of the first author's last name followed by a number used for filing publications in the user's database. Thus, a paper by A. Michels with a number of 6454 containing densities would be labeled MICH6454.PVT or MICH6454.TPV. The data should be entered into the file with temperature first, followed by pressure and/or density, and finally the property that was measured. The values should be entered as given in the manuscript without doing any preliminary unit conversions. In some cases, offsets or other factors may be needed for correcting the data–this can be done by adding simple operators to each value, for example: "3.935+0.01" or "853.33*1.02+0.1". No spaces should be included and the operation is performed from left to right.

An input file should then be created with an extension of INP. The first line of the input file contains the type of data, which can be one of the following:

```
TPV         Temperature, pressure, density (or specific volume)
TP          Temperature, pressure (used for mixture VLE)
Cv          Temperature, density, isochoric heat capacity
TPCv        Temperature, pressure, isochoric heat capacity
Cp          Temperature, pressure, isobaric heat capacity
TDCp        Temperature, density, isobaric heat capacity
Csat        Temperature, saturation heat capacity
W           Temperature, pressure, speed of sound
TDW         Temperature, density, speed of sound
ITC         Temperature, pressure, isothermal throttling coefficient
B           Temperature, second virial coefficient
C           Temperature, third virial coefficient
Cp0         Temperature, ideal gas isobaric heat capacity
U           Temperature, density, internal energy
H           Temperature, pressure, enthalpy
HVap        Temperature, heat of vaporization
JT          Temperature, pressure, Joule-Thomson coefficient
Vis         Temperature, pressure, viscosity
TDVis       Temperature, density, viscosity
Tcx         Temperature, pressure, thermal conductivity
TDTcx       Temperature, density, thermal conductivity
ST          Temperature, surface tension
Pv          Temperature, vapor pressure
Pl          Temperature, liquid pressure (for mixtures)
Ps          Temperature, sublimation pressure
Pm          Temperature, melting pressure
Dl          Temperature, saturated liquid density
Dv          Temperature, saturated vapor density
```

The second line contains the units to which the data will be converted. For TPV data, this would be:

```
"K90" "MPa" "mol/dm^3"
```

The third line contains the molecular weight of the fluid. For a mixture, a –1 should be entered on this line and the molar masses will be determined from the fluid names on the "@TYPE=" line. For VLE mixture data, a fourth line should be added with the characters "@2PHASE" to indicate that both liquid and vapor compositions are included in the file. For mixtures, an additional line should be added to describe which fluids are contained in a mixture. For a methane/ethane mixture, this would read "@TYPE=C1C2".

Additionally, the number of fluids in the mixture should be added in the first line, i.e., "TPV-2" for a binary, or "TPV-3" for a ternary.

The subsequent lines in the file list the names of the files that contain experimental data and their units, like this:

```
6454 "MICHELS " "MICH6454.CH4" "K48" "ATM" "KG/M^3"
```

There are four sections to this line: the filing number for the publication, the first 8 letters of the first author's last name, the file name containing the experimental data, and the units of the data in the file. The units of the data are listed in three parts: temperature, pressure or density, and the property measured. The units for all three must be listed; for data that have only two columns, the third unit should be entered as "". A number of files can be converted at the same time by listing all files of the same data type in the input file.

For compositions entered as volume percent, the flag "VL" can be appended after the units in the input file. On the "@VOLUME=" line, enter all the molar densities of each component at 25 °C and 1 atm, or any other condition at which the volume percents were measured. For compositions entered as weight percent (or fraction), use the flag "WT" after the units in the input file. For equilibrium ratios (K values), enter the flag "K" after the units. In the data file, enter the K values after the temperature and the pressure. To do the conversion to mole fraction, all K values must be available.

The "@AMAGAT=" line is only necessary for a conversion of volumes that are listed as a 'pv' or amagat units. The amagat factor is the density of the fluid in mol/l at 1 atm and 0 °C that can be obtained from REFPROP. If an amagat factor is needed for a mixture, obtain the amagat factor for all of the fluids. List all the factors for the mixture, in order, on the AMAGAT line.

To change units halfway through a file, use "@TUNIT=", "@PUNIT=", or "@DUNIT=". To change fluids halfway through a file, use "@TYPE=". Use the "Stop" statement to terminate the program and ignore all the lines below it.

For mixtures, VLE data contains bubble and dew point pressures and temperatures, along with the associated compositions of the liquid and vapor phases. If a publication includes the density of the saturated liquid or vapor, then each measurement should be entered twice, once in the VLE directory (without the density, just the pressure and temperature), and then in the PVT directory as if it were a single phase point. If a publication include only the density and temperature of the saturation state, then a -1 can be entered for the pressure if it is a liquid state, and a -2 can be entered for the pressure if it is a vapor state.

When the input file is complete, use the program CONVERT.EXE found in the SOURCEB directory to convert the data. Add the name of the input file to the command line used to call the program. (In PM, type AVC while highlighting the INP file.) This will create a *.RAW file that contains all of the data converted to a common set of units and ready for fitting. The numbers will all be located in the correct columns and it will place the proper formatting that the minimize program requires. This program should be

run either from the command line while in the appropriate directory, or from a batch file that is set up to run the conversions. Including the option "/x" after the path will run the conversion without further user intervention.

If a conversion will not execute properly, check to be sure that the listed files in the INP file are identical to those on your hard drive. Use the "Stop" statement to terminate the program, thus allowing you to enter comments below this line. Use the "Skip" statement to stop the program before converting confidential or unwanted data. Use "x" to temporarily quit the program, in which case a message will appear to remind you that there are additional data below the "x". This is useful when checking one dataset at a time.

With the *.RAW files created, the data in these files can be combined into one file that will be used in fitting. An editor can be used to edit the PVT file first and then include the other files (such as CP, SND, etc.) at the bottom of the file. The lines with the formatting should be deleted (they start with an open parenthesis and contains strings such as "F10.6,F10.4"). Alternatively, the *.ID file can be created by issuing the DOS command: "Copy *.RAW R161.ID" where the asterisk is used as shown in this line, not replaced by the file names. (In PM, type C to issue a DOS command.)

In the *.ID file, the numbers and characters have to be aligned within certain columns. The numbers start at the beginning of each row and consist of 10 spaces each. Thus, the pressure usually comes first, located in columns 1-10. The densities are located in columns 11-20 and the temperatures in columns 21-30. The author's last name generally starts in column 52, although there are some exemptions to this. These exemptions are handled automatically by the CONVERT program. In most cases, each line ends in column 80, thus the data point number falls in columns 77-80. During the fitting process, you will weight individual data points and will put the weight of the point starting in column 82. (To view the data in PM, type in A/E/Enter to activate the KEDIT program.)

Each data point needs to be labeled for the fitter to identify its property type. For some of these (DL, DV, and PV), the conversion program will automatically add the data type starting in column 66. For others, it will add a line with an @ symbol and the property type before the beginning of the data, and you should copy the property type on this line into the 11 spaces available starting in column 66. For the property type TPV, you should use PVT.

Several data types may need to be changed manually when the data inputs are pressure and temperature, such as with the speed of sound or with isobaric heat capacities. In these cases, the labels "SND" and "CP" are not sufficient enough to indicate the phase of the data. These indicators will cause the fitter to call the saturation routines at the input temperature, and then determine if the phase is liquid or vapor depending on the input pressure and the calculated vapor pressure. This is a very time consuming process and will bring the fitter to a crawl. When this happens, the computer needs to run a search for each property and determine if the density is above or below the critical level, and therefore identifying what form the substance is in. An "L" for liquid or "V" for vapor should be added to the end of the identifier (e.g., "SNDL") to indicate the phase if known.

If you do not know the phase, you can use the EOSPLOT program explained later to calculate the deviations of the data to the equation of state. This program will print out the densities of each data point, from which you can determine the phase of the data (L for densities greater than the critical density and V for densities less than the critical density). Although not recommended, a further small increase in speed can be obtained by adding a density calculated from a preliminary equation to the data point. This density will need to be updated with better equations, and should be removed before doing the final fit. If both PVT and liquid densities are available, fit only the PVT data because iteration is not required to get the deviations. Fitting DL requires a call to the saturation routines, thus drastically slowing down the fitter.

Nonlinear fitting

To run the minimize program, a batch file should be set up to call MINIMIZE.EXE with inputs on the command line to tell which fluid, which data file and which variables are to be fit. (In PM, create a new file by selecting the directory you want the file to be in, and hit E, then type in the name of the file, hit Enter, then hit the spacebar, then press Esc\Q\Enter to exit the file. You will then edit the file by selecting it and hitting AE and then Enter. Add the text below, and then change the R161 to the fluid you will be using.) An example batch file (called MIN.BAT) is given below:

```
REM :RESTART
    D:\REFPROP\FITTER\MINIMIZE R161.FLD R161.ID  100001 000000 000000 100
    D:\REFPROP\FITTER\MINIMIZE R161.FLD R161.ID  100000 000000 000000 -1
    D:\REFPROP\FITTER\MINIMIZE R161.FLD R161.ID  100000 000000 000000 -3
    D:\REFPROP\FITTER\BCD R161.FLD
    GOTO :RESTART
```

The first line should be commented out by adding the letters "REM" in front of the colon until the user is ready to fit the equation multiple times. The inputs on the command lines are first the fluid file, then the file containing the experimental data and constraints, then three sets of numbers to define which variables are to be fit, and then the maximum number of iterations. A negative one in this last spot tells the fitter to only print the deviations, while a value of −3 tells the program to print the effect of each term in the equation as explained later. A "pause" statement can be added to the end of this file if it closes too quickly to see the end results.

The three sets of numbers that are used to define which variables are to be fit each contain six digits, generally either a 0 or 1. The digits are then turned on or off to fit a particular variable. These variables are explained below:

First set:
   First digit: Fit EOS coefficients
   Second digit: Fit Cp0 coefficients (when set to 1)
                      Fit Cp0 coefficients and exponents (when set to 2)
                      Fit Cp0 coefficients, and exponents starting at i=2 (when set to 3)
                      Fit Cp0 coefficients starting at i=2 (when set to 4)
   Third digit: Fit vapor pressure ancillary when set to 1, liquid density ancillary when set to 2, and vapor density ancillary when set to 3

Fourth digit:	Fit critical density when set to 1, critical temperature when set to 2, and both when set to 3
Fifth digit:	Modify all coefficients at once with one term or randomize the coefficients
Sixth digit:	Fit coefficients only

Second set:

First digit:	Fit alpha in critical region terms
Second digit:	Fit beta in critical region terms
Third digit:	Fit gamma in critical region terms
Fourth digit:	Fit delta in critical region terms

Third set:

First digit:	Fit first coefficient in hmx.bnc file
Second digit:	Fit second coefficient in hmx.bnc file
Third digit:	Fit third coefficient in hmx.bnc file
Fourth digit:	Fit fourth coefficient in hmx.bnc file
Fifth digit:	Fit fifth coefficient in hmx.bnc file
Sixth digit:	Fit excess function

When the MINIMIZE program is run, the output on the screen will display the iteration number, then the variable being fit, and then the sum of squares. Additionally, the initial values of the variables will be displayed before fitting begins. The sum of squares will then decrease as the fitting program minimizes the deviations of the equation to the data. When finished, the program will rewrite the *.FLD file with the new coefficients. This file is also occasionally written during the fitting process in case the process is interrupted (either manually or by a power failure). Thus, any time that the fitter appears to be stuck, the program can be stopped (by pressing Ctrl-C or Ctrl-Break) and then restarted.

A file labeled SSQ is created when the program first runs. This file records the latest value of the sum of squares. If for multiple reasons the sum of squares exceeds an old value, the fluid file will not be overwritten with the current coefficients and the message "Bad SSQ" will be displayed. If you have increased the weights in the ID file or added new data points, this will result in a higher SSQ and the error message. In this situation, you should either delete the SSQ file or edit it and increase its value. A higher SSQ can also result if you are using the option to add random errors to the coefficients; in this case, it will not write out the next fld file until it passes below that best value of the sum of squares. If you do not wish the fitter to check for situations with high SSQs, then make the value in the SSQ file negative and the check will be ignored.

To fit exponents in addition to the coefficients, additional files should be created in the directory where the fluid is being fit. These file are called IC, IT, ID, and IL. The first one lists the numbers of the coefficients that are to be fitted. Generally it will contain the numbers 1 through 20 (or how many terms are in the equation) each listed on a separate line. The IT file lists which terms that the fitter should modify the exponent on temperature to decrease the sum of squares. The ID file (not to be confused with the *.ID file that contains all the data) lists the terms for modifying the density exponent, and the IL file lists the terms for modifying the density term inside the $\exp(-\rho^l)$ piece. If a term does not have an $\exp(-\rho^l)$ piece but it is listed in the IL file, then it will be ignored. Also, numbers above the maximum number of terms will be ignored.

The EXPLMTS file contains limits on the coefficients and exponents in the equation and other information used in fitting. An example file look like:

```
1        !check limits
10       !imax
6.       !ti0max        1
1.9      !timax         2
 .3      !timin         3
 .9      !tiemin        4
1.2      !ticmin        5
6.       !ticmax        6
8.       !dimax         7
1.       !dimin         8
6.       !dlimax        9
0.6      !dlimin        10
35.      !almax         11
.6       !almin         12
1550.    !bemax         13
 .4      !bemin         14
1.8      !gamax         15
0.4      !gamin         16
1.8      !demax         17
 .4      !demin         18
-7.      !term2min      30
 00.     !wtsplt
2.       !devexp
0        !iwrite
```

The first line tells whether the limits should be checked or not (i.e., a zero will ignore all of the subsequent lines). The second gives the number of seconds between saving the FLD file. The lines with a min or max on them give the minimum of maximum value that an exponent can have. If the exponent exceeds this value, then a penalty is added to the sum of squares based on the amount the exponent exceeds its limit. The "e" in tiemin indicates the temperature exponent on the terms with the $\exp(-\rho^l)$ piece. The "c" in ticmin indicates the temperature exponent on the critical region terms. The lines labeled almax through demin are the critical region values for alpha, beta, gamma, and delta. The term2min line allows controlling how small the second term can be, thus a -7 would indicate that the absolute value cannot be smaller than 10^{-7}. The line labeled wtsplt is used to force terms apart when the exponents are also being fitted. Quite often two terms will become intercorrelated and forcing them apart will allow the fitter to find new lows. The value given on this line is the weight of the penalty added to the sum of squares. The numbers in the third column indicate where the penalties are listed in the first block of numbers displayed when MINIMIZE.EXE is run.

Fitting experimental data almost always requires making tough decisions about whether data points with high deviations are caused because the experimental data are bad (or have been converted wrong), or because the equation is not working well. If the problem is the equation, then the bad points really are not bad, and you have to leave them in. But many times two data sets will be contrary and you have to pick the best one by finding which data sets go well with each other. For pure fluids, this means finding data sets for both PVT, sound speed, vapor pressures, and heat capacities that all work with one equation. For example, if fitting heat capacity messes up the sound speed, then that means one of them might be bad. Determining which data sets are correct requires

finding those that work well with each other and then making sure that you have not just overfit or badly fit these data.

Fitting equations of state can take many weeks to many months of work. 95% of fitting experimental data within their uncertainties can be done in 5% of the time, but the last 5% of the development of the equation can take 95% of the time. It is this last 5% where the extrapolation behavior is examined and constraints are added (as explained below) to modify a bad surface so that it becomes smooth and well behaved. Many equations of state have been published quickly and carelessly. These equations show good comparisons to the data but almost always show bad behavior in one or more areas, like heat capacities going negative in the liquid phase at temperatures around the triple point resulting in imaginary values for the speed of sound. An equation is NOT finished until the correlator has checked the extrapolation and smoothness of the equation at ALL conditions, including temperatures below the triple point and at temperatures, pressures, and densities at absurdly high conditions. A bad equation will eventually have to be replaced by a good correlator, so take the time to learn about the constraints to produce a good equation.

Fitting mixtures

The minimize.exe program can fit mixture properties as well. A typical batch file would look like this:

```
D:\REFPROP\FITTER\MINIMIZE HMX.BNC ETOH2O.ID      0 0 111101 250 'ETOH2O'
D:\REFPROP\FITTER\MINIMIZE HMX.BNC ETOH2O.ID      0 0 110000  -1 'ETOH2O'
```

This batch file gives the program the mixture coefficient file name (HMX.BNC), the name of the file containing the experimental data (ETOH2O.ID), several sets of digits to specify which parameters are fit, the number of iterations (250 on the first line), and the data type to be fit, which must be enclosed in quotes. In this example ethanol and water are being fit. The −1 on the last line tells the minimizer to not optimize the data and to only printout the deviations between the equation and the data with the current values in the hmx.bnc file. The first two sets of digits on the first line are set to 0. These two were explained above and are used in fitting pure fluids. The third set is explained below:

Third set:
First digit: Fit first coefficient in hmx.bnc file
Second digit: Fit second coefficient in hmx.bnc file
Third digit: Fit third coefficient in hmx.bnc file
Fourth digit: Fit fourth coefficient in hmx.bnc file
Fifth digit: Fit fifth coefficient in hmx.bnc file
Sixth digit: Fit excess function

On the first line of the batch example given above, the first four coefficients are being fit along with the excess function. Generally there is rarely sufficient data to fit the excess function since we have done most of those with good sources of data already. In the case of ethanol/water, we have quite a large amount of data and this new mixture will be added in the next release of Refprop.

To start fitting mixture properties you will need to decide which mixing rule to use. If you scan through the HMX.BNC file you will find there are quite a number of different mixing rules, with LJ6 and KW0 being the most common. These mixing rules are similar and are explained in the following documents:

http://www.gerg.eu/public/uploads/files/publications/technical_monographs/tm15_04.pdf

http://www.boulder.nist.gov/div838/theory/refprop/REF-MIX.PDF

For very limited data, you can obtain the same result from either mixing rule by fitting only the coefficient for the reducing temperature (Equation 6 in the refrigerant paper). For the LJ6 model, this coefficient is the first value after the letters "LJ6" in the HMX.BNC file. Its value can be positive or negative to indicate an upward or downward shift of the reducing line (a zero values indicates no shift). For KW0 this coefficient is the second value after the letters "KW0" (identified as gammaT and is the gamma variable in equation 7.10 of the GERG paper). Its value is 1 for no departure and close to 1 (either higher or lower) for small departures. Since the KW model will slowly become the norm, it is best to start with this one. As an example, the batch file for a simple system would appear as:

```
  D:\REFPROP\FITTER\MINIMIZE HMX.BNC ETOH2O.ID 0 0 010000 250 'C7C8'
```

As the amount of experimental data increases, it becomes more difficult to fit the data using only one parameter unless the mixture is quite ideal. Fitting additional parameters will enable you to achieve a better fit to the data. After fitting only the reducing temperature parameter (gamma_T, 2nd term on the KW line), , you should then fit the reducing parameter for the reduced density (the gamma_v term). This is the 4th number on the KW lines and 2nd on the LJ lines. In fitting this parameter, you should have experimental densities available otherwise you can easily fit VLE data, but badly distort the equation resulting in bad predictions for the density. When a binary mixture becomes even more nonideal, such as the case with argon/carbon dioxide, you can fit the beta terms in the reducing parameters of equations 7.9 and 7.10 of the GERG document (adjusting beta_T before introducing beta_v). These four parameters are often enough to describe many binary mixtures that have limited amounts of data. For systems with enough data to move beyond fitting just the parameters in the reducing equations, you can fit the term F_{ij} that determines how much of the excess function to include in the mixture Helmholtz energy. The F_{ij} term is given in equation 7.12 in the GERG document and in equation 7 in the LJ document. In the KW lines in the HMX.BNC file, this term is the 5th value on the line; in the LJ lines it is the 3rd value on the line. A value of 1 indicates that the full amount of the excess function is added to the mixture Helmholtz energy. A value of 2 indicates that twice the amount is added, and a value of –1 indicates that the excess function is subtracted from the mixture Helmholtz energy. You are not restricted to the sole use of either KW0 or LJ6; you can try other mixing equations that might be similar to your particular situation, such as KW1 for the methane/ethane mixture function, KW3 for the methane/nitrogen mixture function, LJ4 for the R32/R125 mixture function, and so forth. Mixtures of a hydrocarbon with hydrogen might benefit from the use of the KW7 function. Three figures are given below that show why different excess functions are needed for different binary mixtures.
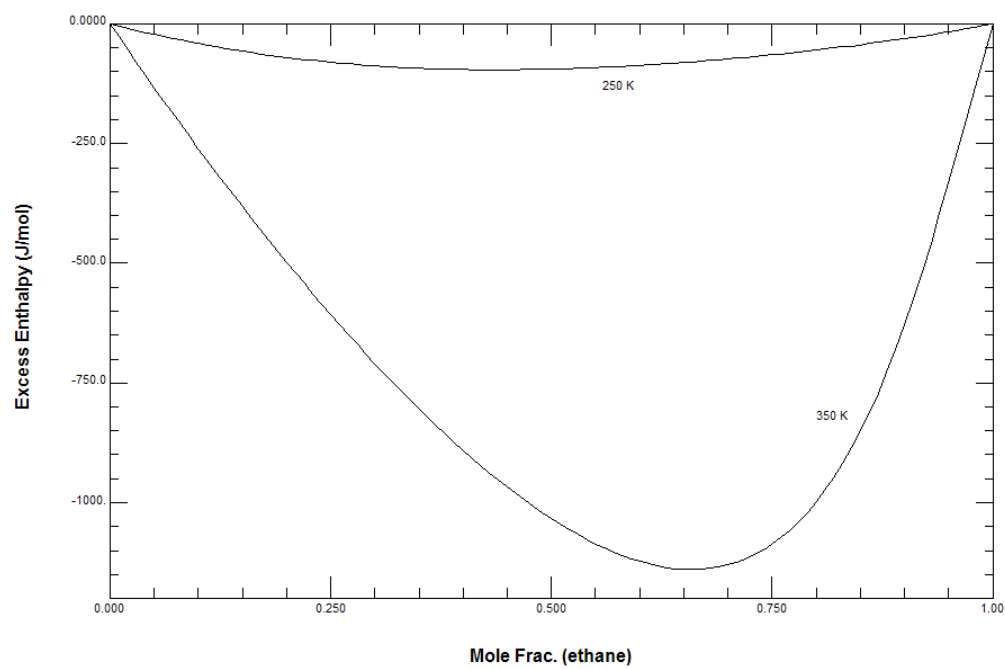
Figure 1. Example of the simple system ethane/butane showing excess enthalpies at 250 K and 350 K and at 10 MPa for ethane/butane. The excess function uses KW0.
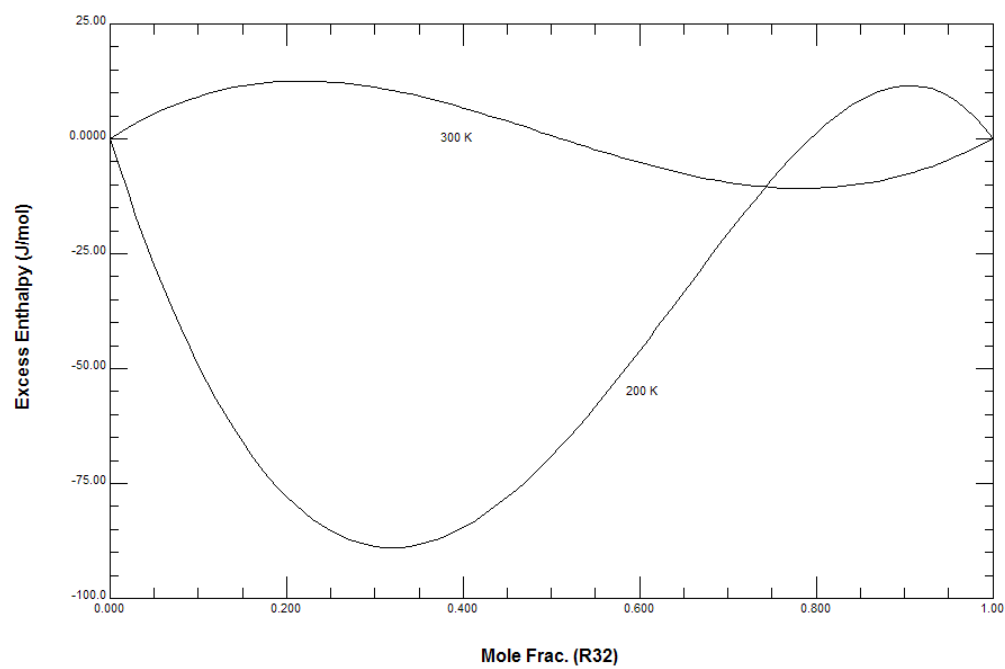
Figure 2. Example of a more complicated system (R32/R125) showing the excess enthalpies at 200 K and 300 K and at 1 MPa. The excess function uses LJ4.
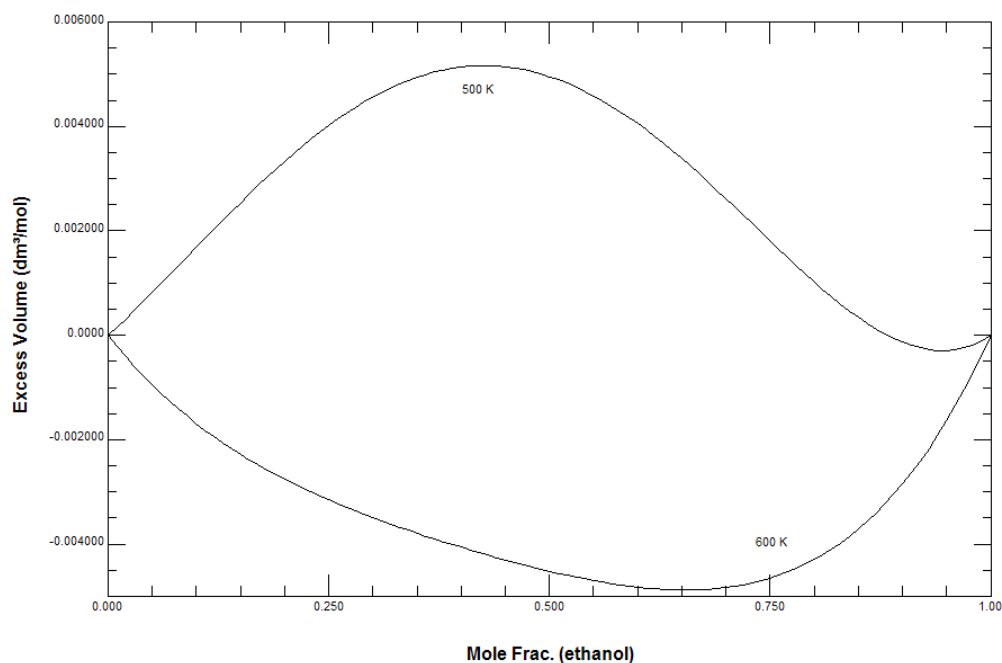
Figure 3. Example of a system containing water (ethanol/water) showing the excess volumes at 500 K and 600 K and at 1 MPa. The excess function uses KW8, a new function developed at NIST using the GERG equations for the reducing values. None of the other functions could mimic the negative excess volume at high mole fractions of ethanol at 500 K.

Before running this you should modify the HMX.BNC file and add the appropriate lines for an heptane/octane mixture. These lines would be similar to the following:

```
?Heptane/Octane            (C7/C8)
?Author name (2009)
142-82-5/111-65-9          KW0   1.    1.    1.    1.    0.    0.
                           TC1   0.    0.    0.    0.    0.    0.
                           VC1   0.    0.    0.    0.    0.    0.
!
```

Duplicate these lines (including the one with the exclamation point), rename the first line of the duplicated lines (which is just a comment, not used by the program) with the appropriate components in the mixture, add your name to the second line, and change the CAS numbers for heptane and octane to that of your desired fluids on the third line. The CAS numbers for the pure fluids can be found in the *.FLD files. Make sure you are not duplicating something like methane/water which already has interaction parameters in it, unless you are sure it is a similar mixture and you want the end results to be similar. The

ideal mixing parameters are all one for the first four numbers when the Kunz-Wagner mixing model is used (KW0). For the Lemmon-Jacobsen mixing model (LJ6), the fourth and fifth parameters are one, and the others are all zero. The lines labeled TC1 and VC1 should contain only zeros. These lines do not contribute to the calculation of thermodynamic properties, except for the simple estimation of the true critical point (which is different from the reducing parameters). The identifier KW0 or LJ6 must start in column 27, but all of the other values are column independent.

The subroutine GETNAMES in the file "OTHER.FOR" in the fitter directory parses the name of the binary mixture. The normal hydrocarbons are specified simply as C1, C8, C16, C24, etc. The isomers are labeled C4I and C5I (isomers above C5 become more complicated). Other gases of interest are labeled as N2, CO2, AR, H2O, H2S, H2, HE, NH3, CO, N2O, SO2, and SF6. The refrigerants used just the number and any identifying letter, for example, 134A152A. In situations without a letter, a dash can be used, thus a mixture of R-22 and R-32 would be "22-32". The not so obvious identifiers are MEO for methanol, ETO for ethanol, EGLY for ethylene glycol, ETY for ethylene, PRY for propylene, TOL for toluene, BEN for benzene, ACT for acetone, CYC or CC6 for cyclohexane, and DME for dimethyl ether.

The ID file contains the experimental data for the mixture. Simply combine the *.RAW files that are generated in the previous sections to form the ID file. The mixture name should be added after the paper number as shown below if it is not already there. In this case, the mixture name is EGLYH2O, indicating data for ethylene glycol with water. Before each data section the type of data should be indicated on a new line starting with the @ symbol. These data types are the same as for pure fluid, with the exception of vapor-liquid equilibrium data, which is labeled as "@VLE". The examples below show that for PVT data, the line contains the pressure, the density, the temperature, the mole fractions of the components, the author name, the paper number, the mixture name, the data point number, and the weight. The compositions are given in the same order as the mixture name. For VLE data, the columns include the pressure, temperature, liquid phase mole fractions of the components, the vapor phase mole fractions, the author name, and so on.

```
@VLE
  0.006053  343.1000 0.8014 0.1986 0.0438 0.9562   CHIAVON166594 EGLYH2O        1
  0.007391  343.1800 0.7581 0.2419 0.0346 0.9654   CHIAVON166594 EGLYH2O        2 1.
  0.008417  343.1200 0.7216 0.2784 0.0271 0.9729   CHIAVON166594 EGLYH2O        3 1.
  0.009926  343.1800 0.6736 0.3264 0.0230 0.9770   CHIAVON166594 EGLYH2O        4 1.
@PVT
  0.101325 55.408557  293.1500 0.0000 1.0000       YANG   166530 EGLYH2O        1
  0.101325 52.177165  293.1500 0.0312 0.9688       YANG   166530 EGLYH2O        2 1.
  0.101325 48.806229  293.1500 0.0676 0.9324       YANG   166530 EGLYH2O        3 1.
@CP
  0.101325            293.1500    75.2858 0.0000 1.0000      YANG   166530 EGLYH2O       1
  0.101325            293.1500    88.0293 0.0312 0.9688      YANG   166530 EGLYH2O       2 .01
  0.101325            293.1500    87.6047 0.0676 0.9324      YANG   166530 EGLYH2O       3
  0.101325            293.1500    86.7040 0.1105 0.8895      YANG   166530 EGLYH2O       4
```

Controlling slopes and curvatures

It is rare that sufficient measurements have been made for a fluid that controlling the shape of the equation is not necessary. More often, insufficient data are available and there are holes where the equation can wiggle such that overfitting other available data

produces what appears to be very good results. The overfitting causes incorrect slopes that give the appearance of a good fit to the limited data but a bad fit in the region where data are not available. To correct this, constraints can be added to the *.ID file. The first and most important constraint forces the critical point to have the correct slope and curvature, that is: $\partial p/\partial \rho$ and $\partial^2 p/\partial \rho^2$ are zero. This constraint appears as:

```
                                                    1 CRITTD       1 1000000.
```

The following two force the second and third virial coefficients to be zero at infinite temperature:

```
                    1.D20      0.                        1 B          2 1.d+2
                    1.D20      0.                        1 C          3 1.d+5
```

The following line would fit the critical pressure at the critical temperature and density (where the values are given for R-161):

```
   5.046      6.28     375.25                CRIT       1 PVT-P       4 1.
```

The data type 'PVT-P' indicates that the deviation in pressure should be added to the sum of squares rather than that for density. The weights at the ends of these lines are arbitrary and can be changed to give the desired fit.

Many other constraints can be added to the fit by adding additional lines to the *.ID file. The data type indicator on these lines consist of three parts, separated by slashes. Examples are given below. The first part indicates what property is being constrained, the second indicates which variables to hold constant and which to vary, and the third indicates what should be constrained: the line, the slope, the acceleration, the third derivative, the fourth derivative, or all of these derivatives. Recognized properties are CV, CP, W, SND, GR, P, DPD, 2PD, Z1, B, C, and D. W and SND both indicate speed of sound. GR is for the Gruneisen parameter, DPD is the first derivate of pressure with respect to density, 2PD is the second, and Z1 is $(Z-1)/\rho$. The second input contains two letters, the first indicates the property being held constant and is either S for saturation, T for temperature, or D for density. The second letter is the property being varied and can be T for temperature, D for density, or L for the logarithm (base 10) of temperature. The third input contains a letter and then either a plus sign, a minus sign, or both. The letter tells what to fit: L for the line, S for the slope, C for the curvature (acceleration), 3 for the third derivative, 4 for the fourth derivative, or A to fit all of the derivatives (but not the line). These indicators can be stacked to do just two or three of them, for example SC would fit the slope and curvature only. The plus or minus sign indicates if the values should be positive or negative. If A was selected, then the signs of the derivatives can be made all positive, all negative, or varied by adding "+−" or "−+" to the third input. The "+−" command would make the slope positive, the curvature negative, the 3rd derivative positive, and the 4th negative. The opposite can be done with the "−+" command. A value of "0" forces the values of the property to be near zero, and the weight can be used to control how close to zero the property should be. Some examples of these are given below:

```
                    10.      200.     10.                    1 C/DT/A+-      3
                    10.      100.     10.                    1 D/DT/L-       2
```

```
         50.      400.       10.                        1 C/DT/4&       5
        600.     2000.       50.                        1 C/DT/A-+      7
       5000.    10000.      200.                        1 B/DT/A-+      7

       -102.     -100.      110.                        1 CV          10
  -1.      5.      200.       5.                         1 CV/ST/A-+    13
  -2.    150.      270.       5.                         1 CV/ST/A+     12
  -1.    340.      364.       1.                         1 CV/ST/A+     11
  -1.    100.      320.       5.                         1 CV/ST/A&     14
  18.     -1.      400.       5.                         1 CV/DT/A-+    18
-100.     5.4     -102.      .005                        1 CV/TD/S+     19
-100.     5.56    -102.     -.005                        1 CV/TD/S-     19

  -1.     60.      200.       5.                         1 GR/ST/A-+    21
   2.     45.      150.       5.                         1 GR/TD/SC+-   22

   1.     24.       40.       4.                         1 P/TD/A+      30
   1.     18.       22.       2.                         1 P/TD/L-      31
-100.      5.        6.      .05                         1 P/TD/S+      32
50000.     1.       50.       1.                         1 2PD/TD/A+    33 1.d-3
 5000.     1.       50.       1.                         1 2PD/TD/A+    34 1.d-3
 -100.  -102.        7.      .05                         1 2PD/TD/S+    37
 -100.     2.5      15.      .2                          1 2PD/TD/40    38
  370.     1.        4.      .1                          1 2PD/TD/30    39

  -1.      1.      100.       1.                         1 CP/ST/A-+    40
  -1.      3.       10.       1.                         1 CP/ST/S0     41
  -1.     50.      200.       5.                         1 W/ST/A-      42
  -1.      6.       10.       1.                         1 W/ST/S-      43
```

In these examples, the first column on the left is the property held constant, the second through fourth columns are the property being varied, with the start in the second column, the end in the third column, and the step in the fourth. When "S" is used for the constant property, a −1 is used for the saturated liquid line and a −2 is used for the saturated vapor line. For any of the properties, a −100 can be used in place of the critical temperature and a −102 can be used in place of the critical density.

Any constraint that calculates properties along the saturation line will require calls to the saturation routines, which is a very time consuming process. The speed of these calculations can be increased dramatically if a vapor pressure ancillary equation is included in the FLD file. In this case the vapor pressure equation is used to calculate an approximate saturation pressure, and the density is then iteratively obtained from the equation of state for either the liquid or the vapor phase. The procedure for adding the ancillary equation is given further on in this document.

Adding weights to data points

The file TOP.ID located at the website below:

http://www.boulder.nist.gov/div838/theory/refprop/TOP.ID

for the fitting files contains the most current and best constraints to add to the top of an ID file. This file is updated as we find better weights and constraints for making equations of state.

The file BLANK.FLD at the website below:

http://www.boulder.nist.gov/div838/theory/refprop/BLANK.FLD

contains a good starting point for any fluid. It was developed using propane data to obtain an equation with the least number of coefficients while still accurately fitting the experimental data. You should rename this file as the name of the fluid with which you are working. The coefficients in this file will give near zero deviations for all of the constraints in the TOP.ID file. Fixed points for a particular fluid must be manually added to this file. These points include the critical temperature, pressure, and density, triple point temperature, molar mass, acentric factor, and EOS limits. The other values such as pressure and density at the triple point and the normal boiling point can be left as is for now, and updated when the fit is done. The EOS limits only need to be approximate at first, and can be set later. Many of the fixed points and other information for fluids can be found at webbook.nist.gov/chemistry.

In addition to the constraints, you should start your fitting process with only one to four PVT or DL data points weighted with very low weights (e.g., "0.0001"). By doing so, the fitter will not become overwhelmed with the high deviations on the data, which could cause it to deviate substantially from the constraints. As the deviations in the data decrease, try increasing the weights on the data by a factor of 10 until you hit "1.". Then start adding additional data (such as vapor pressures, sound speeds, or heat capacities), one point at a time with low weights. As the deviation in each point decreases, increase its weight or slowly add additional points. The idea is to find the best equation with the least amount of data points. Each additional point adds time to the fitting process. The final weight on a data point should reflect the approximate uncertainty. Data points with uncertainties less than 0.08% generally are given final weights of 10 to 100, those with uncertainties of 0.1% to 0.2% are often given final weights of 1, and those with higher uncertainties are given weights of 0.01 to 0.1, depending on how well the correlator wishes to represent those data. High accuracy PVT and sound speeds are typical of the first group, vapor pressures and most PVT data are typical of the second group, and heat capacities and low accuracy speeds of sound are typical of the third group.

To add weights to the *.ID file, edit the file and find the data that you believe are most reliable. At the end of each line (starting in column 82), add a number that represents the weight you wish applied to each data point. Make sure that the ID file was cleaned up by removing any format statements and adding the letters "PVT", "SND", "CV", etc. as explained above at the end of the "Conversion of experimental data" section. Then save the file. To start the fit, run the MIN.BAT file you created earlier. The sum of squares of the deviations will be shown under the column labeled "norm". If the sum of squares is very large, you have either added a weight that is too big, or you have fit a data point that is labeled wrong or may have other problems. Look at the DEV.OUT file to see which point is has a large deviation and either remove that point from the fit or reduce its weight. Once you have the fit running with a good SSQ, remove the "REM" from the first line of the MIN.BAT file to put the fitting into an endless loop.

The purpose of fitting with the MINIMIZE.EXE program is to reduce the sum of squares to as low as possible. As it fits, it will print "Saved at" followed by the sum of squares everytime that the FLD file is rewritten with the new coefficients. If it says "BAD SSQ",

see the notes below about the SSQ file. As it fits, you should continually monitor the DEV.OUT file to see how the deviations for each point are progressing.

Getting unstuck

When an equation has deviated substantially from a good solution, it may be very hard to force it back to a well behaved point. Forcing generally pushes the coefficients in a new direction, but opposite the direction for a good solution. When all attempts to reduce the sum of squares fail, it is often best to return to a previous equation that was well behaved, or to start over with the coefficients taken from the BLANK.FLD file described earlier. As you obtain better and better equations, be sure to save them from time to time with separate file names.

The rectilinear diameter criterion in the ID file is the most common source of getting stuck. You can remove this criterion by adding a # sign at the beginning of the line with RD on it (usually the fist line in the ID file) and rerunning the fit. Once the fitter has moved beyond the initial stuck position, you can add the line back in with a smaller weight initially. The current weight is 1.D8, so try 1.D4 when you add the line back in and slowly increase back to 1.D8.

If this does not work, try removing any other constraint that works along the saturation line, such as the following that include the code "ST".

```
-2.        -95.       -100.      1.                                    1 Z/ST/A-       2 1.d+12
-2.        -65.        -80.      5.                                    1 Z1/ST/A+-     3 1.d+5
-1.        -60.        -80.       5.                                   1 W/ST/A-       4 1.d-2
```

Once the fit comes in, add these lines back in to make sure that your equation meets these constraints.

Another possible scenario for stuck situations is when the fitter finds a local minimum and will not step out of it. The fifth digit in the first set of digits in the MIN.BAT file can be used to add randomness to the coefficients. A value of 3 will add small random numbers to the coefficients, a value of 4 will add a substantial amount, and a value of 5 will cause severe jumps in the coefficients. Generally a value of 4 is good for getting out of local minima, and a value of 5 can be used when watching the fitting process carefully. Both can send the fit into bad areas from which it will not recover, and generally should not be left in this situation overnight. A value of 3 should be used for long periods of unattended fitting. Be sure to save your starting point before using these techniques, as you will often need to return to that point.

Change all gammas and deltas to 1.

Plotting

The *.RAW files can be converted into files used in plotting by running the MAKEREF.EXE program (which is included in the BIBLIO directory, in PM, type A/B/M while on a *.RAW file). Run the MAKEREF program with the name of the *.RAW file as input to the program. When the startup screen appears, at the bottom on

the left the "File contains data" option should be selected. Press F2 to select what kind of data are included in the file. On the top/right side of the screen, select "Graphics" for the output format. Press F1 to run and a *.DEV file will be created. To generate a plot that has the deviations of the data in the DEV file, run the EOSPLOT program (in PM, press A/R/E while on a DEV file). To plot the deviations, use the GRAPHDEV program explained later (in PM, press A/G/A while on a PLT file).

To generate files that can be used to plot the deviations, a batch file (called PLOT.BAT) should be set up calling the EOSPLOT.EXE program (located in the SOURCEB directory) with the names of the files containing the experimental data. An example file is shown below:

```
D:\REFPROP\FITTER\EOSPLOT DEV\R161_B.DEV
D:\REFPROP\FITTER\EOSPLOT DEV\R161CP.DEV
D:\REFPROP\FITTER\EOSPLOT DEV\R161CP0.DEV
D:\REFPROP\FITTER\EOSPLOT DEV\R161DL.DEV
D:\REFPROP\FITTER\EOSPLOT DEV\R161PV.DEV
D:\REFPROP\FITTER\EOSPLOT DEV\R161PVT.DEV
```

The *.DEV files are all located in a DEV subdirectory, and the output of EOSPLOT will be a *.PLT file located in this same directory. The PLT file can be plotted by calling the GRAPHDEV.EXE program with the name of the PLT file on the command line.

The GRAPHDEV program will only allow you to view the plots, not send them to a printer. The printing function is something that will be added later. Printing the PLT files requires the CATSG program…

Equation of State Validating

Once each fit is done, you can copy your fluid file to the c:\program files\refprop\fluids directory before running Refprop. If you are fitting mixture properties, be sure to copy the hmx.bnc file to this directory. If you are updating an equation of state that already exists in Refprop, you may want to rename the old fluid file before copying the new one to this directory.

The TESTPROP program can be used for a number of calculations. Option 1/7 calculates the acentric factor, triple point pressure, triple point density, and normal boiling point temperature. These are all required in the FLD file, although only the acentric factor is used by the program (to get initial estimates for the vapor pressure calculations). Option 3/2 prints out the values of the Cp0 equation to ensure that the equation is behaving properly. These should be printed from 1 K to 10,000 K or higher to make sure that the equation does not go negative values, or to infinity. Option 4/3 calculates the critical point of the equation and compares it to the reducing parameters. For example, for propane it prints out the following:

```
      Reference              EOS              EOS/Ref %
      ---------              -----              ---
t   369.8900000000000 369.8900088996945 100.0000024060381
p   4.251200000000000 4.251165324038544 99.99918432533271
d   5.000000000000000 5.000000628835311 100.0000125767062
```

The first column of numbers gives the values from the FLD file, the second is the location on the equation of state where the first and second derivatives are zero, and the third is the ratio. For the critical temperature and density, these should be between 99.99999 and 100.00000. The critical pressure does not need to be so precise, but should still be very close. It is very important to get the critical temperature very exact. Option 6/5 prints out the ideal curves. These data should be plotted to make sure that the equation is extremely smooth at high state points. For a good example of what these equations should look like, see the high quality equations of Wagner or Lemmon (such as that for propane.)

The program BCD.EXE will eventually be combined into the TESTPROP program. It calculates the 2nd, 3rd, 4th, and 5th virial coefficients of the equation of state. These should be plotted to check their behavior. The R-125 paper of Lemmon et al. (available on the website given previously) gives good examples of what B, C, and the ideal curves should look like. All of the virial coefficients should go to negative infinity at zero K, then should go positive at normal or high temperatures, and then should fall back down to 0 K at extremely high temperatures without ever going negative again.

The Refprop program should be used to make every plot available under the PLOT option. Each one should be checked for smoothness in slope and curvature and to ensure that the equation behaves well. To do this, the limits of the equation should be modified by changing the lower temperature limit (including the triple point temperature) to about ½ of the triple point temperature of the fluid, and by changing the upper three limits to very high values. The R-125 paper gives several examples at extreme conditions. Additional plots or checks should include the following *(additional work and figures are needed on these explanations)*. If you are unsure of what the plot should look like, use the plots of propane as an example.

1. Compare the values of the reducing parameters against those obtained with the equation of state (where the first and second derivatives of pressure with respect to density are zero). The calculated and actual critical temperature should be less than 0.001 K or better, the critical pressure should be within 0.01%, and the critical density within 0.1%. For high accuracy equations of state, the values of 0.000001 K in T_c, 0.001% in P_c, and 0.00001% in ρ_c should be used.

2. Check w vs. T plotting isochores at high values of density

3. Check rectilinear diameter

4. Plot d vs. d2pdd2 with T going from 100 to 5000 (or more), and check values at high d2pdd2

5. Plot log(d) vs. log(d2pdd2)

6. Check d2pdd2 at 5000000 K

7. Check d2pdd2 along the isotherm at Tc. The slope of this line should be constant in the vapor phase to near Dc, with an inflection (but not flat?) at Dc (see ethylene). Look at low densities and make sure the line stays flat, especially at zero density.

8. Check Cv vs. D along Tc to make sure vapor side goes to maximum at Dc

9. Check Cv vs. D along isotherms below Ttrp

10. Check Cv vs. T along sat. line within the last 2 degrees of the critical point to see if Cv goes down at Tc

11. Check Joule-Thomson vs. temperature plot

12. Look at (Z-1)/rho at around 2000 K and make sure there is no curvature at the lower densities

13. Check that csat of the vapor does not go positive for small molecules

14. Check waring criteria (#166260)

15. All default plots in Refprop, especially P-h

16. T-D plot VERY close to critical

17. B-T at low temperature - should continue to go negative

18. Joule curves

19. P-d plots in two phase, and at very high values

20. C should have a maximum around Tc.

21. Check shapes of B, C, D at low and high T

22. Check that there is no curvature in (Z-1)/rho vs. rho at VERY low densities

23. B, C vs. T in vapor phase at low temperature, should be straight lines

24. Check A(two phase)>A(tangent)

25. T vs. Cv-Cp0 plot with isochores, paying special attention for downward curvature in vapor isochores near saturation

26. Plot Cv-Cp0 vs. T by setting #terms for cp0 to 0, then look at isochores

27. Cp-T plotting only isochores. Check for downward curvature at low temps.

28. Cp-T plotting only large values for the isochores (dc*4 to about dc*20). Check for downward curvature

Additional resources

In the SOURCEB directory there are several other programs available that might be of interest to some. These include:

ADDTOINP.EXE	Adds the important information from a data file into the *.INP file. If the units are located on the right hand side of the first three lines (following the data from the paper), these units are incorporated into the INP file. The biblio program is called to get the first author's last name.
ALIGNPRD.EXE	Aligns the periods in a file under those contained on the first line.
REN_EXT.EXE	Renames the extension of a file.
REPEAT.EXE	Brings up a nifty little screen to add commonly used commands. Enter the name of a file on the command line that will store the inputs for the next time the program is run.
ROTATE.EXE	Rotates an HPGL file 90 degrees.
UNITCONV.EXE	Converts values from one unit system to another, including the capability to convert from IPTS-48 and 68 to ITS-90.
NUKEDUPE.EXE	Deletes all the duplicate lines within a file.
SPLIT.EXE	Splits the files up into blocks of common temperatures (???).
SPLTMIXT.EXE	Similar to SPLIT, but for mixtures.
SPLTTEMP.EXE	Splits the files up into blocks of common temperatures.
SPLTTYPE.EXE	Splits the files up into blocks of common data types.
COUNTPNT.EXE	Counts the number of data points in a file and organizes it by author. If STATS has been run first, then the program reads the output of the STATS program and combines it with the output from this program. The end result can be placed directly into a document for publication.
STATS.EXE	Program used to calculate the statistics from the *.PLT files.
STATS2.EXE	Old Stats program

Fitting ancillary equations

The program MINIMIZE.EXE is also used to fit ancillary equations for vapor pressure, saturated liquid density, and saturated vapor density. Ancillary equations are not used by the equation of state for direct calculation of properties, but they can be used as initial estimates for solving the Maxwell criteria. The vapor pressure ancillary is particularly useful for making quick checks to determine if a T,P state point is in the liquid or vapor phase.

Ancillary equations can be generated from either experimental data or from calculations from the equation of state. In the first case, the PV.RAW file generated earlier for use in

fitting the EOS can be used here. In the second case, you can use the TESTPROP program to generate the file used in fitting. Toluene will be used here as an example to show the steps involved with the second case. In the TESTPROP program, select option 2 for ancillary routines and then option 6 to generate the data. This will output a file with lines that look like this:

```
(3F10.6,F10.4,T52,A8,I5,1X,A11,I4,D15.6)
.393935E-7  10.58006 .26618E-7  178.0000          GEN       8012 MAXW           1
.467397E-7  10.56940 .31405E-7  179.0000          GEN       8012 MAXW           2
.553356E-7  10.55876 .36974E-7  180.0000          GEN       8012 MAXW           3
.653730E-7  10.54813 .43440E-7  181.0000          GEN       8012 MAXW           4
.770695E-7  10.53752 .50930E-7  182.0000          GEN       8012 MAXW           5
.906721E-7  10.52693 .59592E-7  183.0000          GEN       8012 MAXW           6
.106460E-6  10.51635 .69588E-7  184.0000          GEN       8012 MAXW           7
.124750E-6  10.50579 .81102E-7  185.0000          GEN       8012 MAXW           8
.145896E-6  10.49524 .94340E-7  186.0000          GEN       8012 MAXW           9
```

Call this file ANC.ID and then add weights to the lines that will be used in the fit. These weights go at the end of the line, like this:

```
(3F10.6,F10.4,T52,A8,I5,1X,A11,I4,D15.6)
.393935E-7  10.58006 .26618E-7  178.0000          GEN       8012 MAXW           1 0.5
.467397E-7  10.56940 .31405E-7  179.0000          GEN       8012 MAXW           2 0.8
.553356E-7  10.55876 .36974E-7  180.0000          GEN       8012 MAXW           3 1.0
.653730E-7  10.54813 .43440E-7  181.0000          GEN       8012 MAXW           4 1.0
```

You will need to add a set of lines to your FLD file as a starting point. The lines below, taken from the propane fluid file, should be added near the bottom of your FLD file, but before the END statement.

```
#PS             !vapor pressure equation
PS5   vapor pressure equation.
?LITERATURE REFERENCE \
?\
!end of info section
0.0                     !lower temperature limit [K]
369.89                  !upper temperature limit [K]
0.0                     !(dummy) upper pressure limit
0.0                     !(dummy) maximum density
369.89   4251.2         !reducing parameters
5 0 0 0 0 0             !number of terms in equation
-6.7722        1.0   !coefficients and exponents
 1.6938        1.5
-1.3341        2.2
-3.1876        4.8
 0.94937       6.2
```

The second line containing "PS5" refers to the type of equation that will be fit. Equation types 1 and 2 use the format

$$\frac{p_\sigma}{p_c} = \sum_{i=1}^{n} n_i \theta^{k_i} \ ,$$

(1)

equation types 3 and 4 use the format

$$\ln\left(\frac{p_\sigma}{p_c}\right) = \sum_{i=1}^{n} n_i \theta^{k_i} \ ,$$

(2)

and equation types 5 and 6 use the format

$$\ln\left(\frac{p_\sigma}{p_c}\right) = \frac{T_c}{T}\sum_{i=1}^{n} n_i\theta^{k_i} \, ,$$

(3)

where $\theta=1-T/T_c$ for equation types 1, 3, and 5, $\theta=(1-T/T_c)^{0.5}$ for vapor pressures of equation types 2, 4, and 6, and $\theta=(1-T/T_c)^{1/3}$ for saturation densities of equation types 2, 4, and 6.

Replace the upper temperature limit and the reducing parameters with the equation of state critical temperature and pressure. Create a batch file like the one below to fit the data:

```
\REFPROP\FITTER\MINIMIZE toluene.fld anc.ID 001000 000000 000000 100
```

The number 001000 tells the program to fit vapor pressure. The number 100 in this line tells the program to execute 100 iterations. Running this program will then fit the coefficients of the vapor pressure equation, write them to the FLD file, and write the deviations between the data and equation in the dev.out file. The 4th column in the file is the percent deviation, and the 5th column is the percent deviation times the weight used in fitting. You can graph this file with the GRAPHDEV program described earlier. In order to decrease the deviations, you can fit the exponents of the equation; however, the exponents 1.0 and 1.5 in the first two terms should not be modified and the others should be greater than 1.5. The technique used to fit exponents is the same as fitting equations of state – create a file called IC and put the numbers 1 to 5 on separate lines (more if you are fitting more than 5 coefficients). Then create a file called IT and put the numbers 3 to 5 on separate lines. Leave the numbers 1 and 2 out since you don't want to fit them.

As in fitting equations of state, the fit will have to be executed multiple times until the sum of squares will no longer reduce. In the toluene example, this procedure reduced the maximum deviation from 0.2% in vapor pressure to 0.01% (0.03% near the critical point), resulting in this fit:

```
5 0 0 0 0 0          !number of terms in equation
-0.741032371962D+01    1.0000000
 0.168156833123D+01    1.5000000
-0.250248318613D+01    2.7954473
-0.248233502099D+01    4.8957906
-0.113105357902D+02   21.1253001
```

Next the exponents should be rounded off and the equation refit. Do this one term at a time, starting with the largest. For example, the 21.1253 number could easily be rounded to 21 (and probably even 20). Remove the number "5" from the IT file and rerun the fit so that the 5th exponent is not changed. Continue this until you have rounded the 4th and then 3rd exponents as well. The coefficients of the final fit should also be rounded (to five or six digits), and cleaned up, as shown below.

```
#PS          !vapor pressure equation
PS5  vapor pressure equation of Lemmon (2009).
```

```
?LITERATURE REFERENCE \
?\
!end of info section
0.0                     !lower temperature limit [K]
591.75                  !upper temperature limit [K]
0.0                     !(dummy) upper pressure limit
0.0                     !(dummy) maximum density
591.75  4126.3          !reducing parameters
5 0 0 0 0 0             !number of terms in equation
-7.4097   1.0           !coefficients and exponents
 1.6785   1.5
-2.5090   2.8
-2.4729   4.9
-10.892  21.0
```

As a final check to make sure you have a good fit, rerun the TESTPROP program, go to option 2, and then run option 14. This will print the differences between the ancillary equation and the values generated from Maxwell solution of the EOS.

Fitting the saturated liquid and vapor densities is similar, except that the values 1 to 5 should be placed in the IT file since it is acceptable to fit all of the exponents. In the batch file, change 001000 to 002000 for liquid density and to 003000 for vapor density.

If you are doing this after the FLD file has been created for the EOS and distributed to others, be sure to add a line similar to the following near the top of the file (but change the initials EWL to your own):

```
! 03-12-09 EWL, add vapor pressure ancillary equation
```

Publishing your equation

…

Once the document has been reviewed by the journal and you have made all applicable changes, a final review must be done by yourself to ensure there are no mistakes in the document. This requires extreme scrutiny so as to not damage your name with mistakes and to avoid having to write an errata. The following should be done:

1. The first pass through the document is to check the English usage, including the following guidelines. The correct use of "that" and "which". No comma is used before "that", and a comma is always used before "which" in situations where either of these words is applicable. The word "data" implies a plural verb: "The data are given…".

2. A second pass checks that the reference numbers in the main part match exactly with those in the references. Reread the document (do not do this simultaneously with #1, you will miss things) and find each reference footnote. Then compare with the references to make sure it aligns and that the word "et al." is used after the first author name when three or more are authors are listed in the publication. If a year is given in the main text or a table, make sure it matches the reference.

3. Pass through the document again looking for the use of an equation number, table number, or figure number, and make sure that they point to the correct equation, table, or figure.

4. If a symbols section is given, make sure that every symbol listed is in the main text, and those in the main text are in the symbols list. Make sure that all physical quantities [variables (T, C_v) or constants (R)] are italicized, and other letters are not (a^r, T_c).

5. Finally, the MOST important item: You must reprogram every equation in your document without the use of any other knowledge. Pretend you are a new engineer and have just read this paper. Now use only the information in the paper to program up the equations in Basic, Fortran, Excel, or whatever and make sure you can obtain the correct answers. If you do not do this, the chances are high that you will have a mistake and people will write you for 20 years because they can not figure out why their program does not work. One of the biggest problems is the signs on the variables inside the Gaussian bell shaped terms.